Computational investigation of BMAA and its carbamate adducts as potential GluR2 modulators

Isidora Diakogiannaki[1,2,⊥], Michail Papadourakis[1,3,⊥], Vasileia Spyridaki[1,4], Zoe Cournia[1,], Andreas Koutselos,[2,*]*

[1]Biomedical Research Foundation, Academy of Athens, 4 Soranou Ephessiou, 11527 Athens, Greece

[2]Department of Chemistry, Physical Chemistry Laboratory, National and Kapodistrian University of Athens, Panepistimiopolis, 15771 Athens, Greece

[3]Department of Nursing, Faculty of Health Sciences, Hellenic Mediterranean University, Heraklion, 71004 Crete, Greece

[4]School of Chemical Engineering, National Technical University of Athens, Heroon Polytechniou 9, 15780 Zografou, Greece

[⊥]These authors contributed equally to this work

*corresponding authors

ABSTRACT

Beta-N-methylamino-L-alanine (BMAA) is a potential neurotoxic non-protein amino acid, which can reach the human body through the food chain. When BMAA interacts with bicarbonate in the human body, carbamate adducts are produced, which share high structural similarity with the neurotransmitter glutamate. It is believed that the L-carbamate adducts of BMAA bind in the glutamate binding site of glutamate receptor 2 type AMPA iGluR (GluR2) causing neurodegenerative effects of diseases such as Alzheimer's and Parkinson's diseases. However, the mechanism of BMAA action and its carbamate adducts bound to GluR2 has not been yet elucidated. Here, we investigate the binding modes and the affinity of BMAA, and its carbamate adducts to GluR2 in comparison to the natural agonist, glutamate, to understand whether these can act as GluR2 modulators. Initially, we perform Molecular Dynamics (MD) simulations of BMAA and its carbamate adducts bound to GluR2 to examine the stability of the ligands in the S1/S2 ligand-binding domain of the receptor. In addition, we perform alchemical free energy calculations to compute the difference in the free energy of binding of the beta-carbamate adduct of BMAA to GluR2 compared to glutamate. Our findings indicate that carbamate adducts of BMAA, and glutamate remain stable in the binding site of the GluR2 compared to BMAA. Additionally, alchemical free energy results reveal that glutamate and the beta-carbamate adduct of BMAA have comparable binding affinity in the GluR2. These results provide a rationale that BMAA carbamate adducts may be in fact the modulators of GluR2 and not BMAA itself.

INTRODUCTION

Beta-N-methylamino-L-alanine (BMAA) is a natural, neurotoxic non-protein amino acid that is produced by a range of ecologically diverse phytoplankton groups such as cyanobacteria, diatoms and dinoflagellates.[1] Specifically, the seeds of cyanobacteria that live in the roots of the cycad tree are a common food source included in flour or eaten by animals consumed by humans. Thus, BMAA may end up in humans through the food chain. In addition, BMAA may reach the human body through the food chain in any aquatic system as BMAA is transferred from cyanobacteria via zooplankton to organisms at higher trophic levels.[2] It has been established that BMAA has neurotoxic and neuro-excitatory properties and that is a source of neurodegenerative disorders in humans; specifically, BMAA has been detected in postmortem brain and spinal cord tissues of Amyotrophic Lateral Sclerosis (ALS), Parkinson's and Alzheimer's patients.[3]

Neurotoxicity of BMAA is dependent on the presence of bicarbonate, which is produced from the interaction of carbon dioxide with water.[3] BMAA is non-toxic in a physiological salt solution, but in presence of bicarbonate that is present in nature and the human body, a carbamylation reaction takes place that produces BMAA carbamate adducts such as alpha-carbamate and beta-carbamate in a ratio of 86:14, respectively.[3] BMAA and its carbamate adducts share a high structural similarity with the natural agonist glutamate (Figure 1).[4] Glutamate and structurally similar substances, such as quinoxalinedione derivatives, activate several glutamate receptors (GluRs), which act as neurotransmitters in the nervous system.[5] GluRs are divided into two subcategories, the metabotropic receptors (mGluR) and the ionotropic receptors (iGluR) that include kainate, α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) and N-methyl-D-aspartate

(NMDA).[6,7] The activation of iGluRs is important in the development and function of the nervous system, while they are also essential in memory and learning. Dysfunction of iGluRs leads to excitotoxic cell death.[6] It has been therefore hypothesized that BMAA's role in the onset and progression of neurodegenerative diseases might be linked to the dysfunction of iGluRs, as BMAA and its carbamate adducts might bind in the binding site of iGluRs, due to their structural resemblance with glutamate.[8,9]

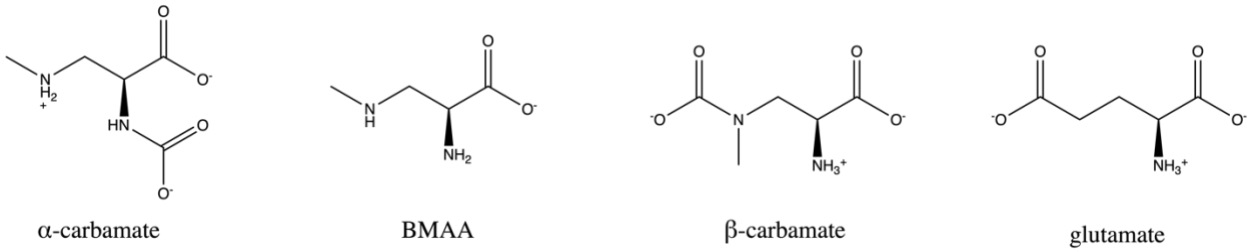

Figure 1. Structures of alpha-carbamate, BMAA, beta-carbamate (L-isomers) and the natural agonist glutamate.

A well-studied subtype of AMPA receptors that may be influenced from the binding of BMAA and its carbamate adducts is the glutamate receptor 2 type AMPA iGluR (GluR2).[10] The binding of glutamate to the GluR2 has been examined extensively both with experimental and computational methods. Crystal structures of glutamate bound to GluR2 (PDB ID: 3DP6[11], 1FTJ[12]) reveal that the binding pocket for glutamate is situated between the two domains, S1 and S2, of the ligand binding core.[6,12,13] The glutamate free energy of binding to the GluR2 based on competitive binding experiments with ^{3}H-AMPA has been determined to be -8.4 kcal/mol.[12] In addition, umbrella sampling and microsecond-long Molecular Dynamics (MD) simulations have been successfully employed to characterize the energetics of the glutamate binding to the GluR2

ligand-binding domain (LBD). It has been estimated that the binding of the natural agonist and the formation of a closed clamshell monomer of GluR2 LBD occur with the release of ca. 4 kcal/mol energy per monomer.[14] Furthermore, atomistic MD simulations have been used to probe glutamate binding in the GluR2.[15] Two possible pathways with which glutamate traverses during ligand binding have been identified using the string method, and the overall free energy difference between the initial unbound state (protein in the apo state, PDB: 1FTO[12]) and the final bound state in pathways 1 and 2 has been estimated to be -5.8 kcal/mol and -8.8 kcal/mol, respectively. The pathway with the lowest free energy of binding corresponds to the glutamate binding pose with PDB ID: 1FTJ, which is the structure that we also used in our RBFE calculations herein. Finally, the standard binding free energy of five ligands, including glutamate, to the GluR2 has been calculated using free energy perturbation calculations (FEP) coupled with MD simulations.[16] The standard binding free energy for glutamate to S1/S2 GluR2 using PDB ID: 1FTJ has been estimated to be -8.5 ± 1.8 kcal/mol, in agreement with experimental results from competitive binding experiments with ^{3}H-AMPA (-8.4 kcal/mol).[16] Based on the aforementioned studies, we confirmed that PDB: 1FTJ is a suitable starting structure for MD/alchemical free energy calculations with glutamate as a ligand.

While studies suggest that the toxic activity of BMAA is linked to the binding of BMAA and its carbamate adducts to the GluR2 glutamate binding site, to the best of our knowledge, there are no previous computational or experimental studies on the binding of these molecules to GluR2 and their mechanism of binding and energetics have not been yet elucidated[8,9]. Therefore, we set out to investigate the stability and the binding affinity of BMAA and its carbamate adducts (L-isomers) to GluR2 in comparison to the natural agonist, glutamate, to assess whether these molecules could bind on GluR2. If their affinity to GluR2 is similar or stronger to glutamate affinity, BMAA and

its carbamate adducts could potentially provoke dysfunction in neurons, leading to neurodegenerative diseases through binding to the GluR2. Herein, we study the stability of BMAA and its adducts in the GluR2 using atomistic MD simulations and relative binding free energy (RBFE) calculations using two force fields to quantify the binding of the beta-carbamate adduct of BMAA with respect to glutamate. Our findings suggest that BMAA carbamate adducts may be in fact the modulators of GluR2 and not BMAA itself.

METHODS

Ligand Structural Similarity and Selection of Crystal Structure for the Calculations

To examine the 2D structural similarity of BMAA and its carbamate adducts with respect to the natural agonist glutamate, the Tanimoto coefficient[17] (Tc) of each ligand compared to glutamate was calculated using ChemBioServer 2.0[18] (see the SI). In addition, we also computed the 3D similarities between the compounds using the Shape Screening module provided by Schrödinger 2020.3 suite[19] and the Pharmacophore type scoring function provided by Phase[20,21].

The selection of the receptor crystal structure has a significant effect on the successful prediction of the stability of a protein-ligand complex[22] and thus, the atomistic model of the glutamate-GluR2 complex was selected based on the following criteria: 1) resolution lower than 3.0 Å, 2) crystallization conditions close to body temperature (T = 310 K) and pH (pH =7.4) (see also the Supporting information). The crystal structure with PDB ID: 1FTJ[12] met the above criteria as it includes the S1/S2 binding core of GluR2 in complex with glutamate and it is crystallized at 277 K, pH = 6.5 and resolution = 1.9 Å. Moreover, it has been successfully used in recovering the energetics of glutamate binding using computational work, in excellent agreement with experiments (see the introduction). As PDB: 1FTJ is from *Rattus norvegicus*, we compared the rat

protein sequence with a human protein sequence using BLAST[23], which showed a 99.5% sequence similarity between the two proteins. The only amino acid change on the S1/S2 domain is a G231R mutation that is located at ca. 17 Å distance from the center of mass of the native agonist, glutamate. We therefore selected 1FTJ[12] to perform the MD and alchemical free energy simulations. Chain A of the GluR2 ligand binding domain was isolated from the full trimer to speed up the simulations (Figure 2). Full details of the model construction and the protein modeling are presented in the SI.

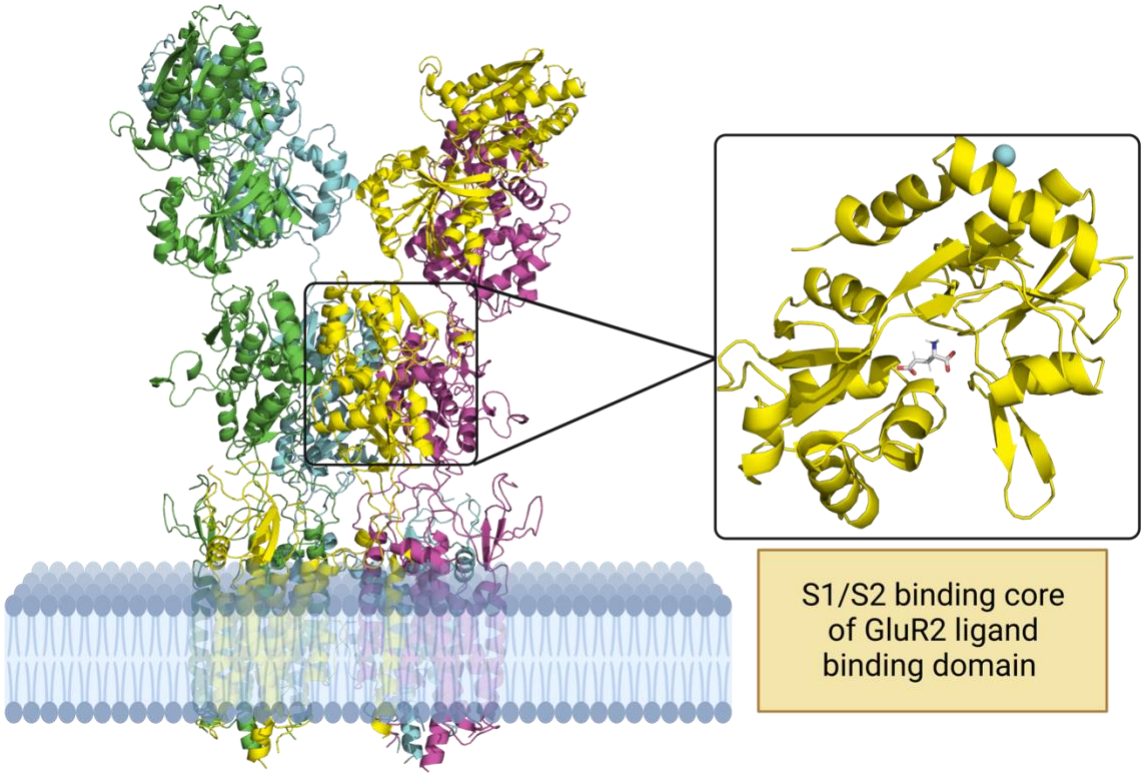

Figure 2. Structure of the GluR2 embedded in a phospholipid bilayer. Inset: chain A of the S1/S2 binding core of GluR2 ligand binding domain of PDB ID: 1FTJ is illustrated together with the bound native agonist glutamate shown in stick representation.

Docking

To create starting configurations for BMAA and its carbamate adducts binding on GluR2, we first assessed the ability of Glide 6.7[24–26] to correctly dock glutamate in the binding pocket of GluR2. We performed self-docking of glutamate in the binding pocket of GluR2 using default parameters. (Figure S1). BMAA, alpha-carbamate and beta-carbamate were also docked into GluR2 using the default parameters of Glide 6.7 to generate initial positions for the MD simulations. The reported binding free energies from the docking calculations were calculated using the Glide SP scoring function, provided by the Glide 6.7 algorithm.

*MD Simulations with NAMD2.14/*OPLS-AA/M/CM1A

Three replicas for each system, namely glutamate/GluR2, BMAA/GluR2, alpha-carbamate/GluR2 and beta-carbamate/GluR2, were simulated for 100 ns using the NAMD 2.14 software. For the replica simulations all conditions were identical except for different initial velocities. For the protein, the OPLS-AA/M[27] force field was used. Ligand parameters were retrieved from the LigParGen[28] webserver (OPLS-AA/CM1A[29] force field) and the TIP3P[30] potential was used for water. The systems were solvated in a cubic box of 15 Å length in each dimension with 45,891 water molecules. Counter ions were added to neutralize the total charge. The long-range electrostatic interactions were treated using the particle-mesh Ewald summation method (PME)[31]. A time step of 2 fs was used. The temperature was kept constant at 310 K using the Langevin thermostat[32] with a time constant of 1 ps. The pressure was isotropically maintained at 1 atm using the Nosé-Hoover barostat[33]. The barostat oscillation period was set at 100 fs, and the barostat damping time scale at 50 fs. The non-bonded potential energy functions (electrostatic and van der Waals) were smoothly truncated between cut-off distances of 10-12 Å.

Prior to MD simulations, all structures were relaxed with 50,000 steps of energy minimization using steepest descent. Heating followed, where the temperature was incrementally changed from 10 to 310 K for 0.6 ns. Then, the systems were equilibrated in the NPT ensemble for 1 ns. Finally, unbiased MD simulations were carried out for 100 ns. To assess the stability of the ligands in the binding cavity of the GluR2, the RMSD of each ligand was calculated with respect to the initial frame using *RMSD Trajectory tool* of VMD 1.9.3[34].

MD Simulations with AMBER20/ff19SB/GAFF2

The same procedure for model construction, equilibration, and MD production protocols was followed using the AMBER20[35] software with the ff19SB[36] force field for the protein and GAFF2[37] parameters that use AM1-BCC[38] charges for the ligands. The Monte Carlo barostat[39] of AMBER20 was used and the pressure was set to 1 atm. The number of steps between the performed volume change attempts was equal to 100.

Relative binding free energy calculations

The relative binding free energy of glutamate bound on GluR2 compared to beta-carbamate was computed with alchemical relative free energy calculations[40–42]. An example of the thermodynamic cycle used for this study is illustrated in Figure 3. For more information on alchemical free energy theory, please refer to the Supporting Information or recent reviews.[41]

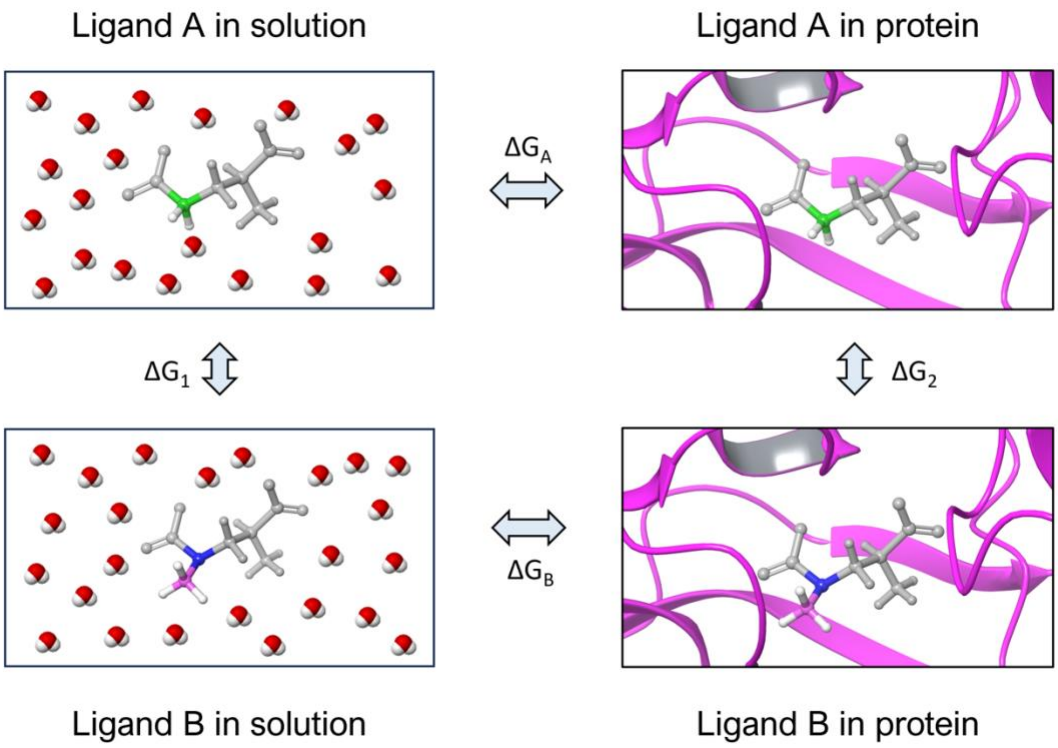

Figure 3. Thermodynamic cycle for the perturbation of ligand A (glutamate) to ligand B (beta-carbamate), in which a methylene group coloured in green is replaced by a nitrogen coloured in blue. The relative binding free energy ($\Delta\Delta G_{binding}$) of beta-carbamate with respect to glutamate is depicted from the horizontal processes (ΔG_B - ΔG_A) and it is equal to the perturbation of glutamate to beta-carbamate in both solution and complex shown by the two vertical processes ($\Delta G_2 - \Delta G_1$). The double arrow for each process illustrates that both forward and backward perturbations are performed.

The perturbation between two ligands, in our case glutamate and beta-carbamate, is typically performed by connecting the two systems using a coupling parameter λ, such that the potential energy function U(λ) interpolates between the initial state (glutamate, λ=0.0) and the final

state (beta-carbamate, λ=1.0). To increase the phase space overlap between the two ligands, we introduced two intermediate molecules, named intermediate-1 and intermediate-2 (Figure 4). Intermediate-1 replaces a carbon-hydrogen bond of the glutamate side chain with nitrogen, and intermediate-2 replaces the nitrogen attached to the methyl group of beta-carbamate with a carbon carbon-hydrogen bond.

The perturbation network for the four compounds was created by the manual connection of the ligands to construct a closed thermodynamic cycle (Figure 4). In the generated relative binding free energy (RBFE) network each edge represents both forward and backward perturbations that are performed in both the bound and the unbound state.

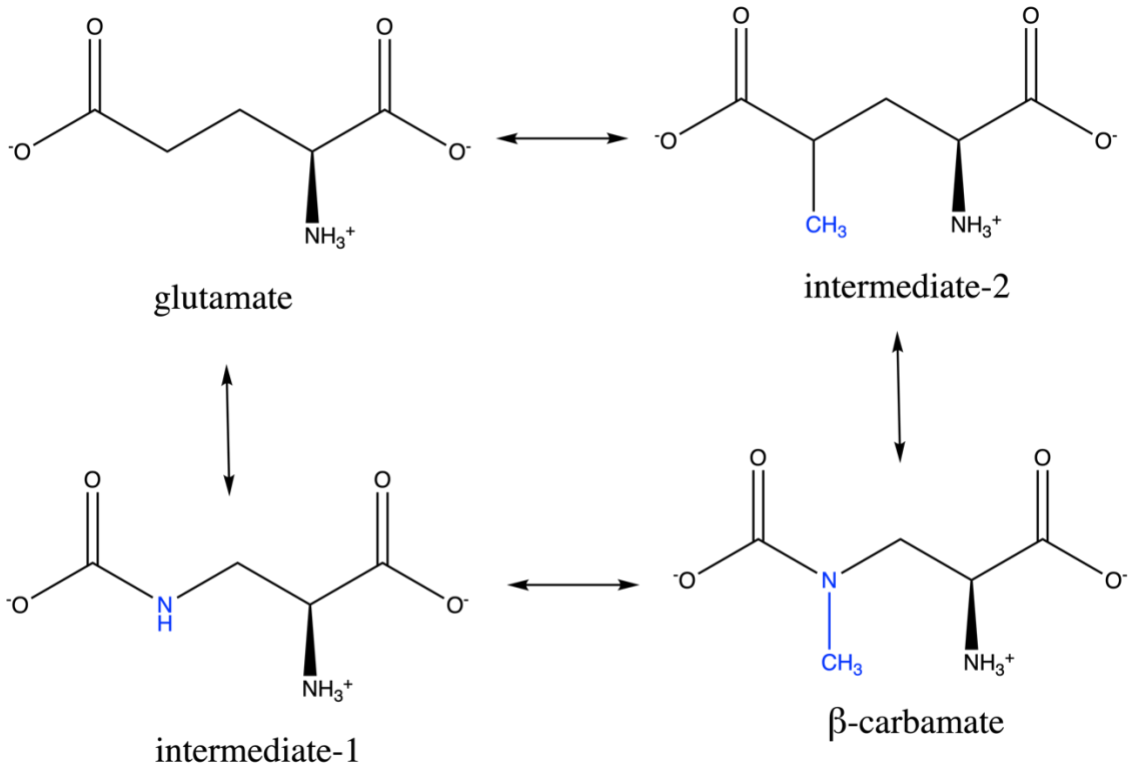

Figure 4. Perturbation network used for the RBFE calculations. Glutamate and beta-carbamate are connected using two intermediate molecules. The perturbed atoms with respect to glutamate are coloured blue.

The primary assumption of RBFE calculations is that all ligands in the congeneric series retain the same binding mode. Therefore, via Maestro Schrödinger Suite 2020-3,[19] beta-carbamate, intermediate-1 and intermediate-2 were overlaid using as a starting structure the crystallographic binding orientation of glutamate into the chain A of GluR2 (Figure S2).

FEP/MD simulations using NAMD 2.14/OPLS-AA

For FEP/MD simulations using NAMD 2.14/OPLS-AA, we generated all input files using the FEPrepare webserver[43], which automates the setup procedure for relative binding FEP calculations for the dual-topology scheme of NAMD using OPLS-AA force field topology and parameter files. Then, all structures were equilibrated in the NVT and NPT ensemble for 1 and 2 ns, respectively. Production run simulations were run for 10 ns with NAMD 2.14 in the NPT ensemble. The number of equidistant λ windows employed for each perturbation was 17 (values for each window ranged between 0.00 - 1.00). A 2 fs timestep was used and a softcore potential was applied to keep pairwise interaction energies finite for all configurations and provide smooth free energy curves for all the simulations[44,45]. A Langevin thermostat and a Nosé-Hoover barostat were applied for the control of temperature and pressure, respectively. The damping coefficient for Langevin dynamics was set at 1 ps, while the barostat oscillation period was set at 100 fs, and the barostat damping time scale at 50 fs. Finally, a 12 Å atom-based cutoff distance for the non-bonded interactions was used and Coulombic interactions were handled using PME.

TI/HREMD simulations using AMBER20/ff14SB

AMBER20 input files were generated using ProFESSA workflow,[46] which automates the setup procedure for the alchemical enhanced sampling (ACES[47]) protocol implemented in the GPU-accelerated AMBER free energy MD engine. In this protocol, a two-state simulation setup

in conjunction with Hamiltonian replica-exchange molecular dynamics (HREMD[48]) was employed to increase the conformational sampling of the alchemical free energy calculations.[49] The equilibration protocol was divided into two phases; the first phase consisted of rigorous equilibration of the initial (λ=0) and end (λ=1) states. First, a minimization with Cartesian restraints relative to the starting structure was implemented to all non-solvent atoms with a force constant of 5 kcal/mol/Å^2 for 5000 steps followed by 5000 steps of minimization of the full system without any restraint. Then, the same restraints were employed for the system to perform an NPT equilibration for 1.2 ns. Subsequently, the system was heated at a fixed volume with 310 K for 1 ns followed by 1 ns equilibration with the NPT ensemble. Then, annealing was performed, where the system was heated to 600 K for 0.1 ns, then stayed with 600 K for another for 0.3 ns and finally cooled down to 310 K in the last 0.4 ns. At that point, the Cartesian restraints were gradually released during a series of simulations of 0.4 ns with the NPT ensemble. Finally, the system was simulated with no restraints with an NPT ensemble for another 0.4 ns.

A second phase of equilibration simulations then followed, where 25 λ states were generated and equilibrated independently. The first half of the λ windows were generated from the equilibrated λ=0 state and the second half of the λ windows were generated from the equilibrated λ=1 phase. The full system was minimized for 5000 steps using steepest descent, then heated at fixed volume with 310 K for 1 ns and equilibrated in the NPT ensemble for 4 ns. Finally, production runs were initiated from the final equilibrated structures, and 10 ns MD simulations were performed for each λ window using a 2 fs timestep. A recently developed robust smoothstep softcore potential was applied to avoid the endpoint catastrophe, particle collapse and large gradient-jump problems often encountered in alchemical free energy simulations[50]. For temperature control at 310 K, a Langevin thermostat was employed with a friction constant of 2.0

ps^{-1}. On the other hand, the pressure was maintained at 1 atm using a Monte Carlo barostat (100 time steps between isotropic box scaling attempts). For the short non-bonded interactions, a 12 Å atom-based cutoff distance was used, while the long-range electrostatics were evaluated with the PME method. The AMBER20/ff14SB/TI protocol was performed three times, using different initial velocities drawn from the Maxwell-Boltzmann distribution.

Analysis of RBFE calculations

To estimate the binding free energy differences from the NAMD2.14/OPLS-AA/FEP simulations, the Multistate Bennett Acceptance Ratio (MBAR[51]) was implemented using the ParseFEP plugin[52] of VMD 1.9.3. Differences in binding free energies from the AMBER20/ff14SB/TI simulations were also estimated with MBAR using python scripts kindly provided by the Darrin York group. The reported binding free energies from the AMBER20/ff14SB/TI protocol are the mean of the three runs, while statistical uncertainties are given as one standard error of the mean.

RESULTS AND DISCUSSION

Ligand Structural Similarity and Generation of Initial Poses

To assess the ability of BMAA and its carbamate adducts to act as GluR2 agonists, we first calculated their 2D structural similarity with respect to glutamate, which revealed a high structural similarity of the molecules under study with the natural agonist with Tc=0.5 (see Methods for a definition of the Tc) for alpha- and beta-carbamates and 0.47 for BMAA (Table S1). Moreover, we performed 3D similarity calculations of the compounds under study using glutamate as the query structure of the Shape Screening calculations. The results showed moderate similarity scores

of beta-carbamate (0.53), alpha-carbamate (0.44) and BMAA (0.50) highlighting that beta-carbamate has the highest binding pose similarity compared to the natural agonist (Table S2).

To test whether Glide can produce valid starting structures for MD simulations, we self-docked glutamate to the GluR2 crystal structure (PDB ID: 1FTJ) with Glide 6.7[24–26]; the RMSD between the docked and the crystal glutamate poses was 1 Å (Figure S1) showing that Glide can accurately predict the binding pose of glutamate.

As starting structures for the MD simulations of BMAA and its carbamate adducts, we used the top docked pose from the Glide docking protocol. However, since BMAA has a single acidic group compared to glutamate, the acidic group could potentially bind in both directions, adopting two flipped binding modes. To investigate the possibility of a flipped BMAA binding pose compared to the top Glide docked pose, we exported the first 13 poses from the Glide docking protocol. From the 13 poses, only the 7th (ΔG = -4.04 kcal/mol) and 8th (ΔG = -3.92 kcal/mol) docking pose presented the acid in the opposite direction, having a ca. 2.39 kcal/mol difference from the top docked pose (ΔG = -6.39 kcal/mol). Therefore, we proceeded with the most energetically favourable docking pose of BMAA, and its carbamate adducts to perform the MD simulations.

MD simulations

We then investigated the binding of BMAA and its carbamate adducts to GluR2, by performing 100 ns equilibrium MD simulations of glutamate, BMAA, alpha-carbamate, and beta-carbamate in complex with the GluR2 using two different protocols: NAMD 2.14[53] software with the OPLS-AA force-field[27–29] (NAMD2.14/OPLS-AA) and AMBER20[35] software using the ff19SB[36] and GAFF2[37] force-fields (AMBER20/ff19SB), all in triplicates with identical starting conditions but different initial velocities. Calculation of the RMSD with respect to the initial frame showed that

all four molecules stay inside the binding pocket of the receptor using NAMD2.14/OPLS-AA protocol ($RMSD_{glutamate}$ = 4.6 ± 0.8 Å, $RMSD_{BMAA}$ = 3.3 ± 1.6 Å, $RMSD_{alpha\text{-}carbamate}$ = 2.8 ± 0.2 Å, $RMSD_{beta\text{-}carbamate}$ = 3.2 ± 0.6 Å) (Figure S3, Table 1) except for BMAA in replica 1 (RMSD = 28.0 ± 18.1 Å) and beta-carbamate in replica 3 (RMSD = 11.4 ± 10.7 Å) that became solvent exposed at 27 ns and 45 ns, respectively. Using AMBER20/ff19SB, glutamate (RMSD = 1.7 ± 0.2 Å), alpha-carbamate (RMSD = 1.7 ± 0.1 Å), and beta-carbamate (RMSD = 2.5 ± 0.1 Å) remained inside the binding site of the GluR2 (Figure S4, Table 2) for all simulations. BMAA drifted off the binding site in two out of three replicas at 74 ns in replica 1 (RMSD = 11.4 ± 15.8 Å) and at 21 ns in replica 3 (RMSD = 30.0 ± 18.4 Å). For the calculation of the average RMSD between the three replicas for both protocols, listed in Table 1 and Table 2, we did not include the RMSD values from the simulations of the solvent-exposed molecules. The average RMSD values obtained from NAMD2.14/OPLS-AA were higher than those observed when using AMBER20/ff19SB with an average difference of ca. 1.50 Å, suggesting that OPLS-AA parameters produce more flexible simulations. Based on our results, we conclude that the substrate glutamate and the ligands alpha-carbamate, and beta-carbamate are stable inside the binding pocket of the GluR2 after 100 ns, while BMAA is less stable and drifts off the binding site of the receptor.

Table 1. Average values of RMSD and standard deviation of each protein-ligand complex were calculated over the 100 ns MD simulations using NAMD2.14/OPLS-AA.

Replica ID	RMSD Glutamate (Å)	RMSD BMAA (Å)	RMSD alpha-carbamate (Å)	RMSD beta-carbamate (Å)
1	5.9 ± 1.4	28.0 ± 18.1	3.1 ± 1.2	2.6 ± 0.3
2	3.3 ± 0.8	1.7 ± 0.3	2.8 ± 0.4	3.8 ± 1.1

3	4.7 ± 1.6	4.8 ± 1.3	2.4 ± 0.7	11.4 ± 10.7
Average RMSD	4.6 ± 1.3	3.3 ± 1.0	2.8 ± 0.8	3.2 ± 0.7

Table 2. Average values of RMSD and standard deviation of each protein-ligand complex were calculated over the 100 ns MD simulations using AMBER20/ff19SB.

Replica ID	RMSD Glutamate (Å)	RMSD BMAA (Å)	RMSD alpha-carbamate (Å)	RMSD beta-carbamate (Å)
1	1.5 ± 0.3	11.4 ± 15.8	1.8 ± 0.3	2.5 ± 1.0
2	2.0 ± 1.0	3.0 ± 0.7	1.9 ± 0.1	2.4 ± 0.8
3	1.6 ± 0.5	30.0 ± 18.4	1.5 ± 0.2	2.6 ± 0.7
Average RMSD	1.7 ± 0.7	-	1.7 ± 0.2	2.5 ± 0.8

To investigate the structural basis of binding of glutamate and beta-carbamate with GluR2, which is the most structurally similar to glutamate (Tc = 0.50 for 2D similarity and Tc = 0.53 for 3D similarity) we used the MDTraj[54] python package and calculated the pairwise RMSDs between the trajectory conformations to find the first cluster representative from each trajectory. The most important interactions between glutamate and GluR2, which are present in at least three replicas with both simulation protocols, are with amino acids Pro89, Thr91, Arg96, Ser142 and Glu193 (Figure 5A). Hydrogen bonds are formed between the hydrogen of the amide group and the hydrogen of the hydroxyl group of Ser142 and the side-chain carboxyl oxygen of glutamate. Glutamate also interacts with ion-dipole interactions with Pro89, Thr91 and Glu193. Finally, a salt bridge interaction is present between the positively charged nitrogen of the guanidinium group of Arg96 and the negatively charged backbone carboxyl oxygen of the substrate. Beta-carbamate's most important interactions with GluR2 are also with residues Pro89, Thr91, Arg96, Ser142 and

Glu193 (Figure 5B). Replica 3 from NAMD2.14/OPLS-AA protocol was not taken into consideration, as beta-carbamate drifts off the binding site of the receptor. Beta-carbamate forms ion-dipole interactions with Pro89, Thr91, Ser142 and Glu193. Finally, there is a salt bridge formed between the positively charged nitrogen of the guanidium group of Arg96 and the backbone negatively charged carboxyl oxygen of beta-carbamate. Thus, a comparison of the structural basis of binding to GluR2 between glutamate and beta-carbamate highlights the high similarity of formed interactions between the two ligands. Ion-dipole interactions of glutamate with Thr91 were present across all replicas for both protocols compared to the corresponding beta-carbamate/Thr91 interactions apart from NAMD2.14/OPLS-AA replica2 simulation. The salt-bridge interaction of beta-carbamate with Arg96 was conserved on all the MD simulations, whereas glutamate/Arg96 interaction was not formed on NAMD2.14/OPLS-AA replica2 simulation.

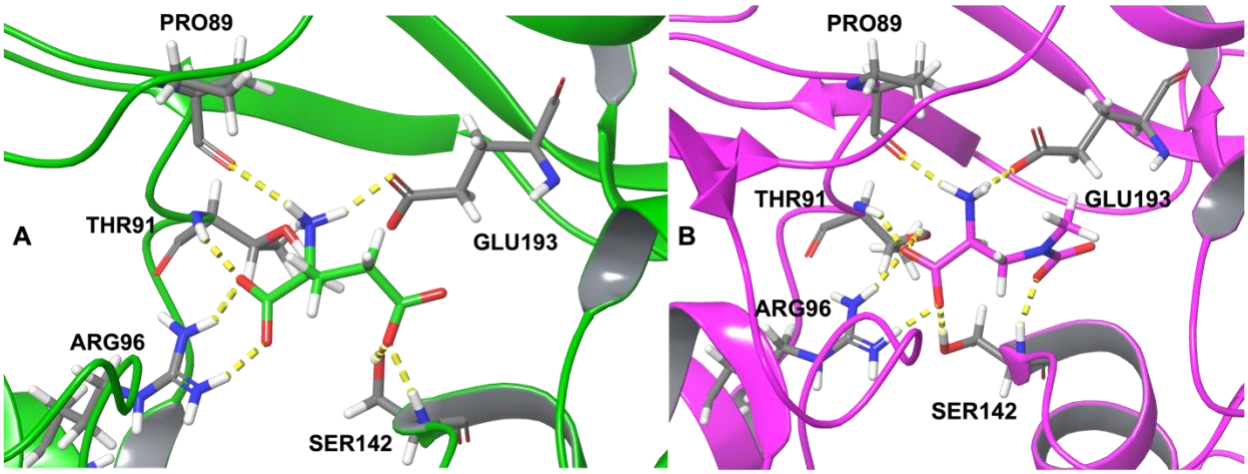

Figure 5. Interactions between A) glutamate and GluR2 and B) beta-carbamate and GluR2.

Relative RBFE calculations

The relative binding free energy of beta-carbamate with respect to the natural agonist glutamate in the GluR2 was determined by alchemical free energy calculations (Figure 3). We examined beta-carbamate because it is the most structurally similar to glutamate (Tc = 0.50 for 2D similarity and Tc = 0.53 for 3D similarity) and remained inside the binding site of GluR2 during the MD simulations. Prior to the RBFE calculations, we examined whether the beta-carbamate could adopt multiple binding modes within the pocket of GluR2 using Glide docking calculations. We exported 16 docking poses and observed that the 3rd (ΔG = -7.35 kcal/mol), 8th (ΔG = -5.75 kcal/mol), and 10th (ΔG = -5.20 kcal/mol) docking poses flip with respect to the glutamate PDB ID 1FTJ pose. The flipped binding pose with the most favourable free energy had a 1.27 kcal/mol difference from the dominant binding pose (ΔG = -8.62 kcal/mol). Therefore, flipping of beta-carbamate during the course of the RBFE calculations should be carefully considered. We thus performed two different protocols: FEP coupled with MD simulations using the NAMD 2.14[53,55] software and the OPLS-AA[27–29] force-field (NAMD2.14/OPLS-AA/FEP) and Thermodynamic Integration (TI) coupled with MD simulation using the AMBER20[35,56] software, the ff14SB[57] force-field for the protein, and the GAFF2[37] force-field for the ligands (AMBER20/ff14SB/TI). The latter protocol uses HREMD that accounts for flipping the molecules during the calculations. The ACES protocol implemented in the GPU-accelerated AMBER free energy MD engine has been successfully tested in complex protein-ligand systems that involve ring flips (Cdk2) and concerted ligand and protein side-chain conformational changes (T4-lysozyme).[47] Thus, AMBER20/ff14SB/TI method is able to account for a potential flipped binding mode of the beta-carbamate adduct. In addition, AMBER TI RBFE calculations have also been successfully performed in relative charged ligand perturbations. In the framework of the 2018 Drug Design

Data Resource (D3R) grand challenge 4, all the predicted relative binding free energies of the Cathepsin S dataset were within 2 kcal/mol error compared to the experimental results.[58] The authors concluded that the high accuracy of the predictions resulted from the force field parameters (ff14SB and GAFF) that are also employed in the AMBER20/ff14SB/TI protocol. In addition, RBFE calculations using the same AMBER force fields and TI calculations were also performed in the PTP1B dataset and showed an overall Mean Unsigned Error (MUE) of 1.6 kcal/mol.[59] This is of critical importance as the binding energetics for glutamate as well as carbamate adducts appear to be primarily driven by electrostatics and atomistic force fields tend to overly reward charge interactions.[60] Our AMBER20/ff14SB/TI protocol was repeated two more times to ensure reproducibility of the results.

We initially introduced one intermediate molecule to connect glutamate and beta-carbamate adduct of BMAA to obtain a more accurate estimate of the binding free energy. The perturbation paths formed a closed cycle between these calculated molecules that was used to monitor the convergence of our RBFE calculations (Figure S5). For a converged calculation, the summation of the free energies over the closed cycle should be close to zero. However, when tested with NAMD2.14/OPLS-AA/FEP, the cycle closure error was equal to -2.11 ± 0.27 kcal/mol. The phase space overlap between the λ states are listed in Table S3 First, to overcome this issue, we introduced additional λ windows for each perturbation, where the convergence between neighbouring λ windows was clearly low (overlap < 40%). This increased the average phase space overlap of the glutamate-intermediate1, intermediate1-beta carbamate and beta carbamate-glutamate perturbations by 27.13% (from 32.96% in the λ windows with overlap < 40% to 59.71%), 29.50% (from 30.30% in the λ windows with overlap < 40% to 59.85%) and 37.17% (from 21.27% in the λ windows with overlap < 40% to 64.41%), respectively (see also the SI for

each perturbation). However, the cycled closed error only changed by 0.1 kcal/mol. To reduce the cycle closure error and to increase the phase space overlap between the two ligands, we added another intermediate molecule, intermediate-2 (Figure 4). However, we did not gain a clear benefit in increasing the phase space overlap from adding the second intermediate molecule as the phase space overlap improved in some neighbouring λ windows, but deteriorated in others. Particularly, for the beta carbamate to intermediate2 perturbation, we observed a decrease in the phase space overlap in seven neighbouring λ windows (decrease ranging from 2 to 40%), while an increase in the phase space overlap was detected in eight neighbouring λ windows (increase ranging from 3 to 42%). On the contrary, for the intermediate2 to glutamate perturbation, we detected a decrease in the phase space overlap in eight neighbouring λ windows (decrease ranging from 3 to 19%), while an increase in the phase space overlap was observed in seven neighbouring λ windows (increase ranging from 4 to 29%). On the contrary, the cycle closure error improved significantly with the addition of the extra intermediate molecule (intermediate-2), which is 1.10 ± 0.09 kcal/mol, compared to the initial three-state RBFE calculations (-2.11 ± 0.27 kcal/mol). Simple addition of extra lambda windows to increase the phase space overlap between the two ligands only changed the cycle closure by 0.10 kcal/mol. Therefore, monitoring the cycle closure error of the RBFE simulations leads to the conclusion that a second intermediate molecule should be added to increase the convergence of our results. Introducing intermediate states between adjacent λ windows with poor phase space overlap did not improve the cycle closure error of the three-state RBFE calculations. With the addition of the second intermediate molecule, the convergence of the calculations improved significantly as the cycle closure error was equal to 1.10 ± 0.09 kcal/mol. Therefore, we also employed the same scheme using AMBER20/ff14SB/TI protocol resulting in a cycle closure error equal to -0.57 ± 0.41 kcal/mol.

To ensure the free energy convergence of our calculations, we measured the degree of overlap between the probability distributions characterizing the equilibrium ensembles for the forward and the backward calculations. The phase space overlap between the λ states ranged between 17.7% - 90.5 using NAMD2.14/OPLS-AA/FEP (Figures S6, S8, S10, S12, S14, S16, S18, S20, Table S4), and 74.5% - 98.3% using AMBER20/ff14SB/TI except for the transition from the initial state to the next neighboring state ($\lambda_{0\text{->}1}$), which refers to the intermediate-1→beta-carbamate. In this case, the overlap was 21.3% in the solvent and 22.8% in the complex system (Figures S6, S8, S10, S12, S14, S16, S18, S20, Table S5-S7) suggesting that more neighboring λ windows could be added in this transformation.

Free energy (ΔG) versus λ for the NAMD2.14/OPLS-AA/FEP protocol and derivative of the gradient of the potential energy ($<\partial U/\partial\lambda>$) as a function of λ versus λ for the AMBER20/ff14SB/TI protocol were also used to determine regions of high curvature across the intermediate λ windows (Figures S7, S9, S11, S13, S15, S17, S19, S21). These regions of phase space serve as a helpful diagnostic as they denote that more sampling or more dense lambda spacing is required to reduce the error of the free energy estimate. In our simulations, we observe low curvature sets of ΔG or $<\partial U/\partial\lambda>$ sets, with the possible exception of the $\lambda_{0\text{->}1}$ transition belonging to the intermediate-1→beta-carbamate for both solvent and complex phases. Thus, a good practice would be to ensure a denser λ spacing in this area.

Table 3. Binding free energy differences for each perturbation and the cycle closure error using NAMD2.14/OPLS-AA and AMBER20/ff14SB.

	NAMD/OPLS-AA	AMBER/ff14SB
Perturbation	ΔΔG (kcal/mol)	ΔΔG (kcal/mol)
Glutamate → **intermediate-1**	1.04 ± 0.08	-0.30 ± 0.12
Intermediate-1 → **beta-carbamate**	-1.54 ± 0.06	-2.47 ± 0.03
Beta-carbamate → **intermediate-2**	2.36 ± 0.14	0.97 ± 0.35
Intermediate-2 → **glutamate**	-0.75 ± 0.08	1.23 ± 0.19
Cycle closure error	1.10 ± 0.09	-0.57 ± 0.41

Overall, our results indicate that the beta-carbamate adduct of BMAA is more potent compared to the natural agonist: the difference in free energy of binding (ΔΔG) of beta-carbamate with respect to glutamate is equal to -1.61 ± 0.33 kcal/mol using NAMD2.14/OPLS-AA/FEP protocol and -2.77 ± 0.12 kcal/mol using AMBER20/ff14SB/TI protocol (Table 3). Therefore, beta-carbamate could potentially bind to GluR2 more strongly than glutamate, potentially leading to dysfunction in neurons, leading to neurodegenerative diseases. Finally, both protocols showed comparable binding affinities for each perturbation. The bigger differences in binding free energies observed for Glutamate->intermediate 1 (1.34 kcal/mol) and beta-carbamate->intermediate-2 (1.39 kcal/mol) perturbations can be fully attributed to the error of the method, that it has been calculated between 1.5-2 kcal/mol, or from the different settings (e.g., force-fields, sampling technique) applied in the two protocols.

CONCLUSIONS

BMAA has been linked to multiple neurodegenerative diseases, but its disease-promoting mechanism remains unknown. BMAA exhibits neurotoxic activity only in the presence of HCO_3^- ions, which leads to the formation of carbamates that share high structural similarity with the glutamic acid. Thus, it has been hypothesized that BMAA and/or its carbamate adducts contribute to neurodegenerative diseases by binding to the glutamate binding site of GluR2.[3] Although here we do not attempt to discover the disease-promoting mechanism of BMAA, we provide insights into whether BMAA and/or its carbamate adducts are binders of GluR2, which could potentially modulate GluR2 at its glutamate binding site and possibly contribute to neurodegenerative diseases.

While the differences between BMAA and its carbamate adducts may look structurally trivial, it is known in the medicinal chemistry literature that small changes in ligand structure may lead to large changes in ligand-protein affinity. Medicinal chemists have long been familiar with the "magic methyl" effect, which is the dramatic change in affinity that may be observed by adding a single methyl group.[61] Additions of fluorine, nitrogen and other atoms may give rise to the same effect. Because the small changes between glutamate, BMAA and its carbamate adducts cannot be quantified in terms of their affinity for the GluR2 by simply eyeballing the structure, the application of robust and accurate methodologies are needed to calculate the affinity of glutamate compared to its adducts on GluR2, in addition to experimental results to determine these effects.

To investigate whether BMAA and/or its adducts bind on GluR2, we first performed equilibrium MD simulations to study the stability of glutamate, BMAA and its carbamate adducts on the GluR2 in atomic-level detail using two different protocols, NAMD2.14/OPLS-AA and AMBER20/ff19SB. Our findings show that the substrate glutamate and the ligands alpha-

carbamate and beta-carbamate remain stable inside the binding pocket of the GluR2 for 100 ns, while BMAA is less stable and drifts off the receptor in 50% of the simulations. This observation is in agreement with experimental assays showing that BMAA was non-toxic in the absence of bicarbonate.[3] Moreover, glutamate and beta-carbamate share similar interactions in the binding pocket of the GluR2, possibly due to their high structural similarity.

Next, to compute the difference in the free energy of binding between the beta-carbamate adduct of BMAA and glutamate to the GluR2, we performed alchemical free energy calculations with two different protocols, NAMD2.14/OPLS-AA/FEP and AMBER20/ff14SB/TI. We did not investigate the binding of BMAA to GluR2 using RBFE calculations because the total charge of these molecules within the perturbation would not remain conserved. Changing the total system charge could yield accurate results,[62] but there are limited successful examples of charge changes during RBFE calculations in the literature.[41]

Results from the RBFE calculations between the beta-carbamate adduct of BMAA and glutamate binding to GluR2 reveal that using two intermediate molecules instead of one, which was the original setup, significantly improved the closed cycle error of the simulations. The phase space overlap between the different λ states ranged between 17.7% - 90.5% with NAMD2.14/OPLS-AA/FEP and 74.5% - 98.3% using AMBER20/ff14SB/TI indicating that the latter protocol was more efficient for this study. In both cases, the difference in the free energy of binding between glutamate and beta-carbamate was negative (-1.61 ± 0.33 kcal/mol for NAMD2.14/OPLS-AA/FEP protocol and -2.77 ± 0.12 kcal/mol for AMBER20/ff14SB/TI protocol). Although these values are close to the error of the method, they do suggest that beta-carbamate has a comparable, if not stronger, binding affinity with glutamate. Therefore, based on relative binding free energy calculations and studying the interactions between GluR2 and BMAA and its adducts, we provide

a rationale that BMAA carbamate adducts may be in fact the modulators of GluR2 and not BMAA itself.

This study also employs two different MD packages and force-fields to ensure the consistency of the results across the codes. Both MD protocols were able to predict the stability of glutamate, alpha-carbamate, and beta-carbamate inside the binding pocket of the GluR2 compared to BMAA. The average RMSD values obtained from NAMD2.14/OPLS-AA were ca. 1.50 Å higher than those observed when using the AMBER20/ff19SB protocol. In addition, the free energy results obtained from the two protocols were also comparable within the limits of the error of the method (1.16 kcal/mol). However, the AMBER20/ff14SB/TI was more efficient for this study based on the phase space overlap between the λ neighbours compared to the NAMD2.14/OPLS-AA/FEP protocol. In addition, the ProFESSA workflow made the AMBER/TI easier to use compared to NAMD/FEP.

Although alchemical free energy calculations are now routinely used in the pharmaceutical industry in lead optimization cycles, they are still not streamlined with non-commercial software. In this manuscript, we report problems that we faced during the calculations concerning convergence issues and how we solved them (introduction of two intermediate molecules). We conducted a technically sound study using two different alchemical free energy protocols and given the high interest of the pharmaceutical industry in alchemical free energy calculations, we strongly believe that such studies are important to be demonstrated in the current literature. Especially for the AMBER20/ff14SB/TI protocol, we employed a state-of-the-art methodology that was published recently from Darrin York's group[46,47,49], tested it and provided a proof of concept for its use. Finally, we provide the full dataset for this protocol, as well as for all the other MD and RBFE protocols that we applied.

Overall, both MD protocols highlight the stability of the beta-carbamate/GluR2 complex, while both RBFE protocols predict a comparable binding affinity of the beta-carbamate and the natural agonist glutamate. Therefore, the beta-carbamate adduct of BMAA could potentially bind equipotently or more strongly to the GluR2 compared to glutamate and may provoke dysfunction in neurons, leading to neurodegenerative diseases; this hypothesis should be further examined through experimental assays.

SUPPORTING INFORMATION

The Supporting Information is available free of charge.

Additional information and supporting figures and tables describing the ligands' similarity with glutamate, model construction, relative binding free energy calculations' theory and results concerning self-docking, and alchemical free energy calculations starting structures and results. (PDF)

AUTHOR INFORMATION

Corresponding Author

Andreas Koutselos - *Department of Chemistry, Physical Chemistry Laboratory, National and Kapodistrian University of Athens, Panepistimiopolis, 15771 Athens, Greece*

Email: akoutsel@chem.uoa.gr

Zoe Cournia - Biomedical Research Foundation, Academy of Athens, 4 Soranou Ephessiou, 11527 Athens, Greece; https://orcid.org/0000-0001-9287-364X;

Email: zcournia@bioacademy.gr

Authors

Isidora I. Diakogiannaki - *Biomedical Research Foundation, Academy of Athens, 4 Soranou Ephessiou, 11527 Athens, Greece, Department of Chemistry, Physical Chemistry Laboratory, National and Kapodistrian University of Athens, Panepistimiopolis, 15771 Athens, Greece*

Michail Papadourakis - *Biomedical Research Foundation, Academy of Athens, 4 Soranou Ephessiou, 11527 Athens, Greece*

Vasileia Spyridaki - *Biomedical Research Foundation, Academy of Athens, 4 Soranou Ephessiou, 11527 Athens, Greece*

Present Addresses

Isidora I. Diakogiannaki – Department of Chemistry, University of Naples Federico II, 49 Via Domenico Montesano, 0810678102 Naples, Italy

Author Contributions

I.D., M.P. and V.S. performed computational experiments. I.D., M.P., Z.C. and A.K. wrote and revised the manuscript. Z.C. and A.K. conceived the study and supervised the research. All authors have given approval for the final version of the manuscript.

Funding Sources

Z.C., I.D. and M.P. acknowledge computational time granted from the Greek Research & Technology Network (GRNET) in the National HPC facility ARIS under project IDs pr010017/FEPBMAA, pr012008/BMAA-FEP; M.P. acknowledges the Bodossaki Foundation Scholarships program for financially supporting his research. Z.C. acknowledges funding by the European Union's Horizon 2020 research and innovation programme under the grant agreement No. 857645 (project NI4OS-Europe).

Notes

The authors declare no competing financial interest.

Data and Software Availability. The NAMD 2.14 software package is available at https://www.ks.uiuc.edu/Research/namd/ and the AMBER20 software package is available at http://ambermd.org/. FEPrepare webserver is freely accessible at https://feprepare.vi-seem.eu/ and the LigParGen webserver is freely accessible at http://zarbi.chem.yale.edu/ligpargen/. The input/output files generated during this study are available at https://doi.org/10.5281/zenodo.7730768.

ACKNOWLEDGMENT

We would like to thank Anastasia Hiskia for useful discussion about the possible implication of BMAA in GluR2 biology. Thanks are also extended to the Darrin York group and especially Abir

Ganguly, Shi Zhang and Darrin York, for useful discussions and for providing the ProFESSA workflow, analysis scripts and the AMBER20 software license for AMBER/ff14SB/TI and AMBER/ff19SB protocols. Finally, we want to thank the reviewers that helped improving the manuscript.

REFERENCES

(1) Vyas, K. J.; Weiss, J. H. BMAA – an Unusual Cyanobacterial Neurotoxin. *Amyotrophic Lateral Sclerosis* **2009**, *10* (sup2), 50–55. https://doi.org/10.3109/17482960903268742.

(2) Porojan, C.; Mitrovic, S. M.; Yeo, D. C. J.; Furey, A. Overview of the Potent Cyanobacterial Neurotoxin β-Methylamino-L-Alanine (BMAA) and Its Analytical Determination. *Food Additives & Contaminants: Part A* **2016**, *33* (10), 1570–1586. https://doi.org/10.1080/19440049.2016.1217070.

(3) Diaz-parga, P.; Goto, J. J.; Krishnan, V. V. Chemistry and Chemical Equilibrium Dynamics of BMAA and Its Carbamate Adducts. *Neurotox Res* **2018**, *33* (1), 76–86. https://doi.org/10.1007/s12640-017-9801-2.

(4) Myers, T. G.; Nelson, S. D. Neuroactive Carbamate Adducts of Beta-N-Methylamino-L-Alanine and Ethylenediamine. Detection and Quantitation under Physiological Conditions by 13C NMR. *Journal of Biological Chemistry* **1990**, *265* (18), 10193–10195. https://doi.org/10.1016/S0021-9258(18)86928-9.

(5) Catarzi, D.; Colotta, V.; Varano, F. Competitive Gly/NMDA Receptor Antagonists. *Current Topics in Medicinal Chemistry 6* (8), 809–821.

(6) Mayer, M. L. Crystal Structures of the GluR5 and GluR6 Ligand Binding Cores: Molecular Mechanisms Underlying Kainate Receptor Selectivity. *Neuron* **2005**, *45* (4), 539–552. https://doi.org/10.1016/j.neuron.2005.01.031.

(7) Zhu, S.; Gouaux, E. Structure and Symmetry Inform Gating Principles of Ionotropic Glutamate Receptors. *Neuropharmacology* **2017**, *112* (Pt A), 11–15. https://doi.org/10.1016/j.neuropharm.2016.08.034.

(8) Allen, C. N.; Spencer, P. S.; Carpenter, D. O. Beta-N-Methylamino-L-Alanine in the Presence of Bicarbonate Is an Agonist at Non-N-Methyl-D-Aspartate-Type Receptors. *Neuroscience* **1993**, *54* (3), 567–574. https://doi.org/10.1016/0306-4522(93)90228-8.

(9) Goto, J. J.; Koenig, J. H.; Ikeda, K. The Physiological Effect of Ingested β-N-Methylamino-L-Alanine on a Glutamatergic Synapse in an in Vivo Preparation. *Comp Biochem Physiol C Toxicol Pharmacol* **2012**, *156* (3–4), 171–177. https://doi.org/10.1016/j.cbpc.2012.07.004.

(10) Isaac, J. T. R.; Ashby, M. C.; McBain, C. J. The Role of the GluR2 Subunit in AMPA Receptor Function and Synaptic Plasticity. *Neuron* **2007**, *54* (6), 859–871. https://doi.org/10.1016/j.neuron.2007.06.001.

(11) Ahmed, A. H.; Wang, Q.; Sondermann, H.; Oswald, R. E. Structure of the S1S2 Glutamate Binding Domain of GLuR3. *Proteins: Structure, Function, and Bioinformatics* **2009**, *75* (3), 628–637. https://doi.org/10.1002/prot.22274.
(12) Armstrong, N.; Gouaux, E. Mechanisms for Activation and Antagonism of an AMPA-Sensitive Glutamate Receptor: Crystal Structures of the GluR2 Ligand Binding Core. *Neuron* **2000**, *28* (1), 165–181. https://doi.org/10.1016/S0896-6273(00)00094-5.
(13) Furukawa, H.; Singh, S. K.; Mancusso, R.; Gouaux, E. Subunit Arrangement and Function in NMDA Receptors. *Nature* **2005**, *438* (7065), 185–192. https://doi.org/10.1038/nature04089.
(14) Yonkunas, M.; Buddhadev, M.; Flores Canales, J. C.; Kurnikova, M. G. Configurational Preference of the Glutamate Receptor Ligand Binding Domain Dimers. *Biophys J* **2017**, *112* (11), 2291–2300. https://doi.org/10.1016/j.bpj.2017.04.042.
(15) Yu, A.; Lau, A. Y. Energetics of Glutamate Binding to an Ionotropic Glutamate Receptor. *J. Phys. Chem. B* **2017**, *121* (46), 10436–10442. https://doi.org/10.1021/acs.jpcb.7b06862.
(16) Heinzelmann, G.; Chen, P.-C.; Kuyucak, S. Computation of Standard Binding Free Energies of Polar and Charged Ligands to the Glutamate Receptor GluA2. *J. Phys. Chem. B* **2014**, *118* (7), 1813–1824. https://doi.org/10.1021/jp412195m.
(17) Bajusz, D.; Rácz, A.; Héberger, K. Why Is Tanimoto Index an Appropriate Choice for Fingerprint-Based Similarity Calculations? *Journal of Cheminformatics* **2015**, *7* (1), 20. https://doi.org/10.1186/s13321-015-0069-3.
(18) Karatzas, E.; Zamora, J. E.; Athanasiadis, E.; Dellis, D.; Cournia, Z.; Spyrou, G. M. ChemBioServer 2.0: An Advanced Web Server for Filtering, Clustering and Networking of Chemical Compounds Facilitating Both Drug Discovery and Repurposing. *Bioinformatics* **2020**, *36* (8), 2602–2604. https://doi.org/10.1093/bioinformatics/btz976.
(19) Schrödinger Release 2023-1: Maestro, Schrödinger, LLC, New York, NY, 2023.
(20) Dixon, S. L.; Smondyrev, A. M.; Rao, S. N. PHASE: A Novel Approach to Pharmacophore Modeling and 3D Database Searching. *Chemical Biology & Drug Design* **2006**, *67* (5), 370–372. https://doi.org/10.1111/j.1747-0285.2006.00384.x.
(21) Dixon, S. L.; Smondyrev, A. M.; Knoll, E. H.; Rao, S. N.; Shaw, D. E.; Friesner, R. A. PHASE: A New Engine for Pharmacophore Perception, 3D QSAR Model Development, and 3D Database Screening: 1. Methodology and Preliminary Results. *J Comput Aided Mol Des* **2006**, *20* (10), 647–671. https://doi.org/10.1007/s10822-006-9087-6.
(22) Lionta, E.; Spyrou, G.; K. Vassilatis, D.; Cournia, Z. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Current Topics in Medicinal Chemistry* **2014**, *14* (16), 1923–1938.
(23) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *Journal of Molecular Biology* **1990**, *215* (3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
(24) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. *J. Med. Chem.* **2006**, *49* (21), 6177–6196. https://doi.org/10.1021/jm051256o.
(25) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739–1749. https://doi.org/10.1021/jm0306430.
(26) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47* (7), 1750–1759. https://doi.org/10.1021/jm030644s.

(27) Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. *J. Chem. Theory Comput.* **2015**, *11* (7), 3499–3509. https://doi.org/10.1021/acs.jctc.5b00356.
(28) Dodda, L. S.; Cabeza de Vaca, I.; Tirado-Rives, J.; Jorgensen, W. L. LigParGen Web Server: An Automatic OPLS-AA Parameter Generator for Organic Ligands. *Nucleic Acids Research* **2017**, *45* (W1), W331–W336. https://doi.org/10.1093/nar/gkx312.
(29) Jorgensen, W. L.; Tirado-Rives, J. Potential Energy Functions for Atomic-Level Simulations of Water and Organic and Biomolecular Systems. *Proceedings of the National Academy of Sciences* **2005**, *102* (19), 6665–6670. https://doi.org/10.1073/pnas.0408037102.
(30) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935. https://doi.org/10.1063/1.445869.
(31) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092. https://doi.org/10.1063/1.464397.
(32) Davidchack, R. L.; Handel, R.; Tretyakov, M. V. Langevin Thermostat for Rigid Body Dynamics. *J. Chem. Phys.* **2009**, *130* (23), 234101. https://doi.org/10.1063/1.3149788.
(33) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant Pressure Molecular Dynamics Algorithms. *J. Chem. Phys.* **1994**, *101* (5), 4177–4189. https://doi.org/10.1063/1.467468.
(34) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *Journal of Molecular Graphics* **1996**, *14* (1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.
(35) D.A. Case, K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz,R Harris, S. Izadi, S.A. Iz-Mailov, K. Kasavajhala, A. Kovalenko, R. Krasny, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, V. Man, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, A. Onufriev, F. Pan, S. Pantano, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, L. Wilson, R.M. Wolf, X. Wu, Y. Xiong, Y. Xue, D.M. York and P.A. Kollman (2020), AMBER 2020, University of California, San Francisco.
(36) Tian, C.; Kasavajhala, K.; Belfon, K. A. A.; Raguette, L.; Huang, H.; Migues, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **2020**, *16* (1), 528–552. https://doi.org/10.1021/acs.jctc.9b00591.
(37) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *Journal of Computational Chemistry* **2004**, *25* (9), 1157–1174. https://doi.org/10.1002/jcc.20035.
(38) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *Journal of Computational Chemistry* **2002**, *23* (16), 1623–1641. https://doi.org/10.1002/jcc.10128.
(39) Faller, R.; de Pablo, J. J. Constant Pressure Hybrid Molecular Dynamics–Monte Carlo Simulations. *J. Chem. Phys.* **2002**, *116* (1), 55–59. https://doi.org/10.1063/1.1420460.
(40) Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22* (8), 1420–1426. https://doi.org/10.1063/1.1740409.
(41) Cournia, Z.; Allen, B.; Sherman, W. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. *J. Chem. Inf. Model.* **2017**, *57* (12), 2911–2937. https://doi.org/10.1021/acs.jcim.7b00564.
(42) Athanasiou, C.; Vasilakaki, S.; Dellis, D.; Cournia, Z. Using Physics-Based Pose Predictions and Free Energy Perturbation Calculations to Predict Binding Poses and Relative Binding Affinities for FXR

Ligands in the D3R Grand Challenge 2. *J Comput Aided Mol Des* **2018**, *32* (1), 21–44. https://doi.org/10.1007/s10822-017-0075-9.

(43) Zavitsanou, S.; Tsengenes, A.; Papadourakis, M.; Amendola, G.; Chatzigoulas, A.; Dellis, D.; Cosconati, S.; Cournia, Z. FEPrepare: A Web-Based Tool for Automating the Setup of Relative Binding Free Energy Calculations. *J. Chem. Inf. Model.* **2021**, *61* (9), 4131–4138. https://doi.org/10.1021/acs.jcim.1c00215.

(44) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. Avoiding Singularities and Numerical Instabilities in Free Energy Calculations Based on Molecular Simulations. *Chemical Physics Letters* **1994**, *222* (6), 529–539. https://doi.org/10.1016/0009-2614(94)00397-1.

(45) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. Separation-shifted Scaling, a New Scaling Method for Lennard-Jones Interactions in Thermodynamic Integration. *J. Chem. Phys.* **1994**, *100* (12), 9025–9031. https://doi.org/10.1063/1.466707.

(46) Ganguly, A.; Tsai, H.-C.; Fernández-Pendás, M.; Lee, T.-S.; Giese, T. J.; York, D. M. AMBER Drug Discovery Boost Tools: Automated Workflow for Production Free-Energy Simulation Setup and Analysis (ProFESSA). *J. Chem. Inf. Model.* **2022**, *62* (23), 6069–6083. https://doi.org/10.1021/acs.jcim.2c00879.

(47) Lee, T.-S.; Tsai, H.-C.; Ganguly, A.; York, D. M. ACES: Optimized Alchemically Enhanced Sampling. *J. Chem. Theory Comput.* **2023**. https://doi.org/10.1021/acs.jctc.2c00697.

(48) Meng, Y.; Dashti, D. S.; Roitberg, A. E. Computing Alchemical Free Energy Differences with Hamiltonian Replica Exchange Molecular Dynamics (H-REMD) Simulations. *J Chem Theory Comput* **2011**, *7* (9), 2721–2727. https://doi.org/10.1021/ct200153u.

(49) Tsai, H.-C.; Lee, T.-S.; Ganguly, A.; Giese, T. J.; Ebert, M. C.; Labute, P.; Kenneth M. Merz, J.; York, D. M. AMBER Free Energy Tools: A New Framework for the Design of Optimized Alchemical Transformation Pathways. *Journal of Chemical Theory and Computation* **2023**. https://doi.org/10.1021/acs.jctc.2c00725.

(50) Lee, T.-S.; Lin, Z.; Allen, B. K.; Lin, C.; Radak, B. K.; Tao, Y.; Tsai, H.-C.; Sherman, W.; York, D. M. Improved Alchemical Free Energy Calculations with Optimized Smoothstep Softcore Potentials. *J. Chem. Theory Comput.* **2020**, *16* (9), 5512–5525. https://doi.org/10.1021/acs.jctc.0c00237.

(51) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129* (12), 124105. https://doi.org/10.1063/1.2978177.

(52) Liu, P.; Dehez, F.; Cai, W.; Chipot, C. A Toolkit for the Analysis of Free-Energy Perturbation Calculations. *J. Chem. Theory Comput.* **2012**, *8* (8), 2606–2616. https://doi.org/10.1021/ct300242f.

(53) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *Journal of Computational Chemistry* **2005**, *26* (16), 1781–1802. https://doi.org/10.1002/jcc.20289.

(54) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L.-P.; Lane, T. J.; Pande, V. S. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys J* **2015**, *109* (8), 1528–1532. https://doi.org/10.1016/j.bpj.2015.08.015.

(55) Jiang, W.; Chipot, C.; Roux, B. Computing Relative Binding Affinity of Ligands to Receptor: An Effective Hybrid Single-Dual-Topology Free-Energy Perturbation Approach in NAMD. *J. Chem. Inf. Model.* **2019**, *59* (9), 3794–3802. https://doi.org/10.1021/acs.jcim.9b00362.

(56) Lee, T.-S.; Allen, B. K.; Giese, T. J.; Guo, Z.; Li, P.; Lin, C.; McGee, T. D. Jr.; Pearlman, D. A.; Radak, B. K.; Tao, Y.; Tsai, H.-C.; Xu, H.; Sherman, W.; York, D. M. Alchemical Binding Free Energy Calculations in AMBER20: Advances and Best Practices for Drug Discovery. *J. Chem. Inf. Model.* **2020**, *60* (11), 5595–5623. https://doi.org/10.1021/acs.jcim.0c00613.

(57) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.
(58) Zou, J.; Tian, C.; Simmerling, C. Blinded Prediction of Protein-Ligand Binding Affinity Using Amber Thermodynamic Integration for the 2018 D3R Grand Challenge 4. *J Comput Aided Mol Des* **2019**, *33* (12), 1021–1029. https://doi.org/10.1007/s10822-019-00223-x.
(59) Song, L. F.; Lee, T.-S.; Zhu, C.; York, D. M.; Merz, K. M. Jr. Using AMBER18 for Relative Free Energy Calculations. *J. Chem. Inf. Model.* **2019**, *59* (7), 3128–3135. https://doi.org/10.1021/acs.jcim.9b00105.
(60) Ahmed, M. C.; Papaleo, E.; Lindorff-Larsen, K. How Well Do Force Fields Capture the Strength of Salt Bridges in Proteins? *PeerJ* **2018**, *6*, e4967. https://doi.org/10.7717/peerj.4967.
(61) Schönherr, H.; Cernak, T. Profound Methyl Effects in Drug Discovery and a Call for New C–H Methylation Reactions. *Angewandte Chemie International Edition* **2013**, *52* (47), 12256–12267. https://doi.org/10.1002/anie.201303207.
(62) Zhou, R.; Das, P.; Royyuru, A. K. Single Mutation Induced H3N2 Hemagglutinin Antibody Neutralization: A Free Energy Perturbation Study. *J. Phys. Chem. B* **2008**, *112* (49), 15813–15820. https://doi.org/10.1021/jp805529z.